AD_____

Award Number: W81XWH-12-1-0294

TITLE: High-throughput functional validation of progression drivers in lung adenocarcinoma.

PRINCIPAL INVESTIGATOR: Don L. Gibbons, MD, PhD

CONTRACTING ORGANIZATION: MD Anderson Cancer Center, University of Texas Houston, TX 77030

REPORT DATE: September 2013

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
September 2013	Annual Report	1 September 2012 - 31 August 2013

4. TITLE AND SUBTITLE
High-throughput functional validation of progression drivers in lung adenocarcinoma.

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-12-1-0294

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Don L. Gibbons, MD, PhD

email: dlgibbon@mdanderson.org

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
MD Anderson Cancer Center , ` niversity of` Texas
Houston, TX 77030

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Cancer cells are endowed with diverse biological capabilities driven by an ensemble of inherited, somatic and epigenetic aberrations. As we enter the era of personalized medicine, characterization of the cancer genome has begun and will continue to influence diagnostic and therapeutic decisions in the clinic. Genome profiling technologies are generating a compendium of genomic aberrations in major cancer lineages with the goal of identifying the most promising therapeutic targets and diagnostic biomarkers. The output from these technologies is radically transforming cancer science. At the same time, these efforts are revealing the complexity of cancer genomes, which are comprised of causal "driver" aberrations and many more biologically neutral "passengers" that arise through the unstable nature of tumor genomes. While most cancers acquire one or more well-studied, high frequency driver events (e.g., mutations/gene copy number changes in KRAS, TP53, EGFR, MYC, BRAF, etc.), much less is known about the overly abundant low frequency (<5%) aberrations contributing to tumor progression and response to therapeutics. Comprehensive biological assessment of low frequency aberrations is difficult given their large number and the fact that they may either directly drive tumor progression or indirectly influence tumor behavior through modifying activities of other drivers like KRAS. Moreover, distinguishing driver events from passengers is further complicated by the fact that driver events are shaped by the specific biological context of a given cancer, including its tissue type, microenvironment and other host determinants including the immune system. The primary objective of the Early Investigator Synergistic Idea Award held by Dr. Kenneth Scott and myself is to establish a driver prioritization pipeline to functionally evaluate lung cancer genomics data to identify somatic driver aberrations, beyond the handful of well-characterized genes like oncogenic KRAS and EGFR, that contribute to lung cancer progression, invasion, and metastasis. Our study integrates genetically engineered mouse models of lung cancer, genomics data generated by The Cancer Genome Atlas (TCGA) and functional screens to identify drivers of lung cancer progression. We are using these tools to implement a novel, scalable screening infrastructure that permits high-content, gain-of-function screens to accelerate validation of functional somatic aberrations. This work is possible through advances made in our laboratories that include (1) high-throughput, highly accurate modeling of somatic aberrations into our collection of over 32,000 sequence-verified human genes and (2) a novel molecular barcoding approach that facilitates cost-effective detection of driver events following in vitro and in vivo functional screens. Our Specific Aims are as follows: (1) Construction of a lung cancer somatic driver library; (2) Functional screens for drivers of lung cancer metastasis; (3) Clinicopathological prioritization and validation of top candidates.

15. SUBJECT TERMS- none provided

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE		10	USAMRMC
U	U	U	UU		19b. TELEPHONE NUMBER *(include area code)*

TABLE OF CONTENTS

INTRODUCTION

Cancer cells are endowed with diverse biological capabilities driven by an ensemble of inherited, somatic and epigenetic aberrations. As we enter the era of personalized medicine, characterization of the cancer genome has begun and will continue to influence diagnostic and therapeutic decisions in the clinic. Genome profiling technologies are generating a compendium of genomic aberrations in major cancer lineages with the goal of identifying the most promising therapeutic targets and diagnostic biomarkers. The output from these technologies is radically transforming cancer science. At the same time, these efforts are revealing the complexity of cancer genomes, which are comprised of causal "driver" aberrations and many more biologically neutral "passengers" that arise through the unstable nature of tumor genomes. While most cancers acquire one or more well-studied, high frequency driver events (e.g., mutations/gene copy number changes in *KRAS*, *TP53*, *EGFR*, *MYC*, *BRAF,* etc.), much less is known about the overly abundant low frequency (<5%) aberrations contributing to tumor progression and response to therapeutics. Comprehensive biological assessment of low frequency aberrations is difficult given their large number and the fact that they may either directly drive tumor progression or indirectly influence tumor behavior through modifying activities of other drivers like *KRAS*. Moreover, distinguishing driver events from passengers is further complicated by the fact that driver events are shaped by the specific biological context of a given cancer, including its tissue type, microenvironment and other host determinants including the immune system. The primary objective of the *Early Investigator Synergistic Idea Award* held by Dr. Kenneth Scott and myself is to establish a **driver prioritization pipeline** to functionally evaluate lung cancer genomics data to identify somatic driver aberrations, beyond the handful of well-characterized genes like oncogenic *KRAS* and *EGFR*, that contribute to lung cancer progression, invasion, and metastasis. Our study integrates genetically engineered mouse models of lung cancer, genomics data generated by The Cancer Genome Atlas (TCGA) and functional screens to identify drivers of lung cancer progression. We are using these tools to implement a novel, scalable screening infrastructure that permits high-content, gain-of-function screens to accelerate validation of functional somatic aberrations. This work is possible through advances made in our laboratories that include (1) high-throughput, highly accurate modeling of somatic aberrations into our collection of over 32,000 sequence-verified human genes and (2) a novel molecular barcoding approach that facilitates cost-effective detection of driver events following *in vitro* and *in vivo* functional screens. Our Specific Aims are as follows: (1) Construction of a lung cancer somatic driver library; (2) Functional screens for drivers of lung cancer metastasis; (3) Clinicopathological prioritization and validation of top candidates. **Our overall collaborative progress for the first 12 months of the project is described in the year one report submitted by Dr. Kenneth Scott and herein I outline the specific contributions of my group.**

BODY

Based on our approved Statement of Work, our overall goals, timeline and progress for the first 12 months of this project were as follows:

Specific Aim 1: Construction of a lung cancer somatic driver library (proposed completion, months 3-6).

Please see the overall report from Dr. Scott for this section, as the work was conducted in his laboratory.

Specific Aim 2: Functional screens for drivers of lung cancer metastasis.

Subaim 2.1 *In vitro* screens for cell invasion and anoikis resistance (proposed completion, months 6-12)

In this Subaim we propose to perform parallel *in vitro* screens for gene drivers of cell invasion and anoikis resistance using a 96-well format for cells transduced with single candidate genes from the ORF library constructed in Aim 1. We began by performing extensive optimization of all steps required for these screens, including establishing cell plating and screening conditions (pilot assays) for three separate cell lines: the optimization for the 393P and 393LN murine cell lines (described in our application) were performed in my laboratory and a new cell line derived from human bronchial epithelial cells was performed by Dr. Scott's group (HBECs; see description in Aim 3 below).

Cell Invasion: We first conducted a series of pilot assays employing both 96-well and 24-well cell invasion assays used and previously described by us [1]. We assessed performance of the 393P and 393LN cell lines stably transduced with negative (mCherry) and positive (including *ZEB1* and *HOXA1;* [1],[2]) control drivers of cell invasion in several vector backbones (pLenti6 and pDEST51). These initial studies led us to choose the 393P cell line as the primary screening model for the invasion screens. This decision was based on our published observations and pilot experiments demonstrating that 393P cells (1) form primary tumors upon inoculation of immunocompetent syngeneic animals but do not metastasize, thus permitting us to screen for metastasis drivers, (2) have low baseline invasion activity (Fig. 2A), and (3) can be stimulated to invade in the presence of a *bona fide* invasion drivers, *HOXA1* and *ZEB1* (Fig. 2A).

Based on the successful pilot studies described above, we have initiated screening of our ORF library in 393P. Briefly, 393P cells are infected with lentivirus carrying the indicated genes, selected in puromycin for stable expression, and then screened for *in vitro* invasion using the established conditions (Fig. 2) and controls. Generation of the stable cell lines and cryo-banking were performed by Dr. Kundu in my lab and Dr. Wu in the Scott lab working together. Importantly, the cryo-banked cell lines can be used for (1) future functional and biochemical validation assays and (2) *in vivo* screening assays proposed in Subaim 2.2 and as described below.

We initially proposed to complete the *in vitro* assays during the first year; however, we determined that screening would be more efficient if conducted in a "rolling" fashion whereby individual candidate pools (e.g., 20-30 gene sets of the 279 library panel) were entered into both the *in vitro* and *in vivo* screens simultaneously, thus eliminating the costly need to construct the transduced cell line panel on two separate occasions. Screening a subset of the 279 candidates has already revealed multiple genes with robust pro-invasion activity (Fig. 2B), and we are actively screening all 279 candidates, with the expectation to complete the primary screens within the next 3 months (~01/01/2014).

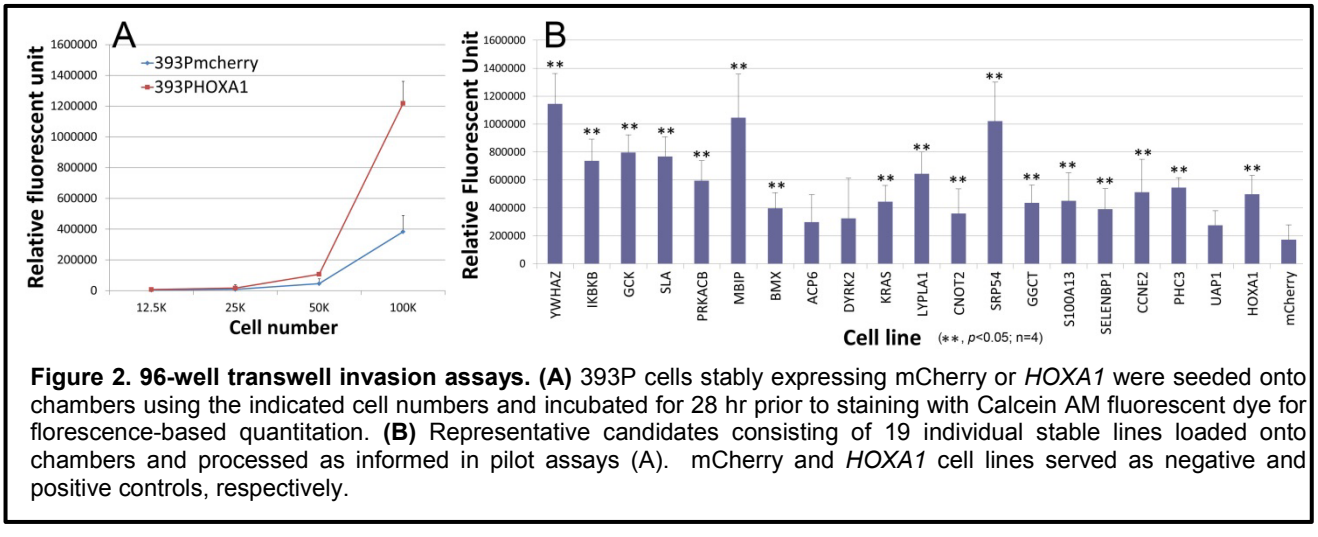

Figure 2. 96-well transwell invasion assays. (A) 393P cells stably expressing mCherry or *HOXA1* were seeded onto chambers using the indicated cell numbers and incubated for 28 hr prior to staining with Calcein AM fluorescent dye for florescence-based quantitation. **(B)** Representative candidates consisting of 19 individual stable lines loaded onto chambers and processed as informed in pilot assays (A). mCherry and *HOXA1* cell lines served as negative and positive controls, respectively.

The top scoring gene candidates from the 96-well screen are subsequently validated by performing invasion and migration assays in the standard 24-well Boyden chamber system in the Gibbons lab. Consistent with the primary screen in 96-well format, we observed significant increases in both migration and invasion for the 393P cells expressing many of the top candidate genes, including *YWHAZ*, *MBIP* and *SRP54*, when compared to the mCherry negative control cells (Fig. 3A, B).

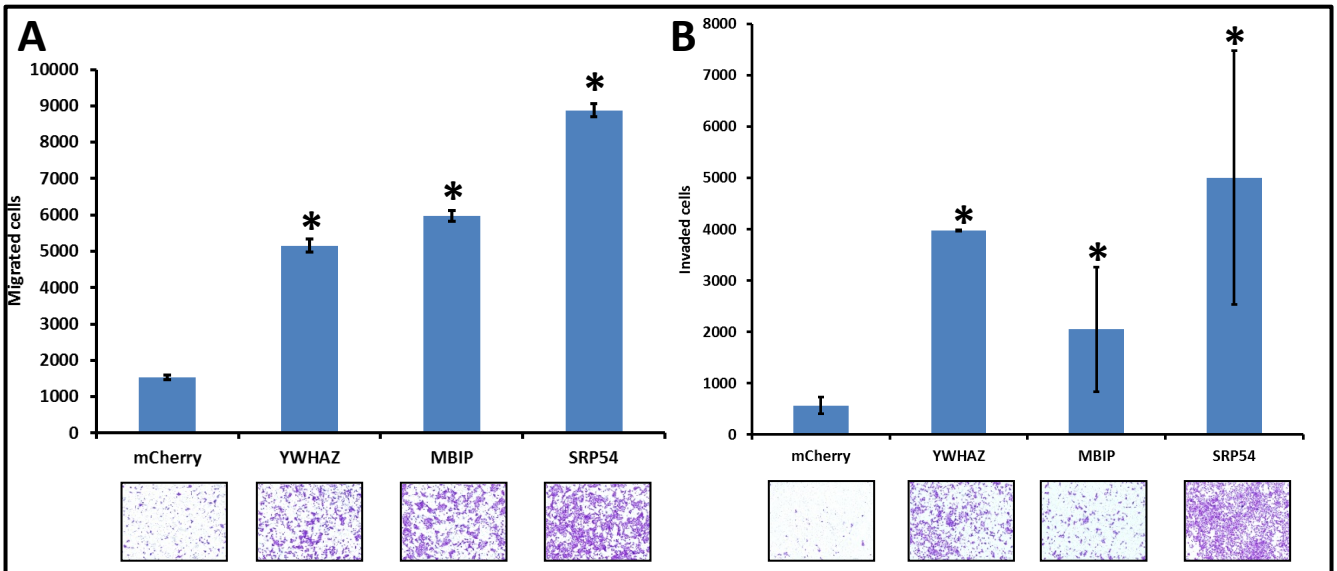

Figure 3. 24-well transwell migration & invasion assays. (A) 393P cells stably expressing mCherry or the indicated genes were seeded onto chambers at 50,000 cells/well, incubated for 16 hr prior to staining with crystal violet and counted using ImageJ software. **(B)** The same stable cell lines were plated onto chambers containing a layer of Matrigel, incubated for 16 hr, stained with crystal violet and counted.

Anoikis Resistance: We have made great efforts to develop the proposed anoikis-resistance assays; however, our efforts have thus far been unsuccessful due to the fact that all cell lines assayed to date (e.g., 393P, 393LN, etc) already exhibit robust anoikis resistance and are therefore not suitable as screening models. Dr. Scott's group has been examining our newly-derived HBEC cell model described below to assess its performance in this assay, and are also currently evaluating use of rat intestinal epithelial (RIE) cells, which is a screenable [3] "generalized" line highly sensitive to anoikis as demonstrated by RIE cells' inability to proliferate under low cell attachment conditions. We anticipate determining whether or not anoikis screens will be feasible using these new cell systems

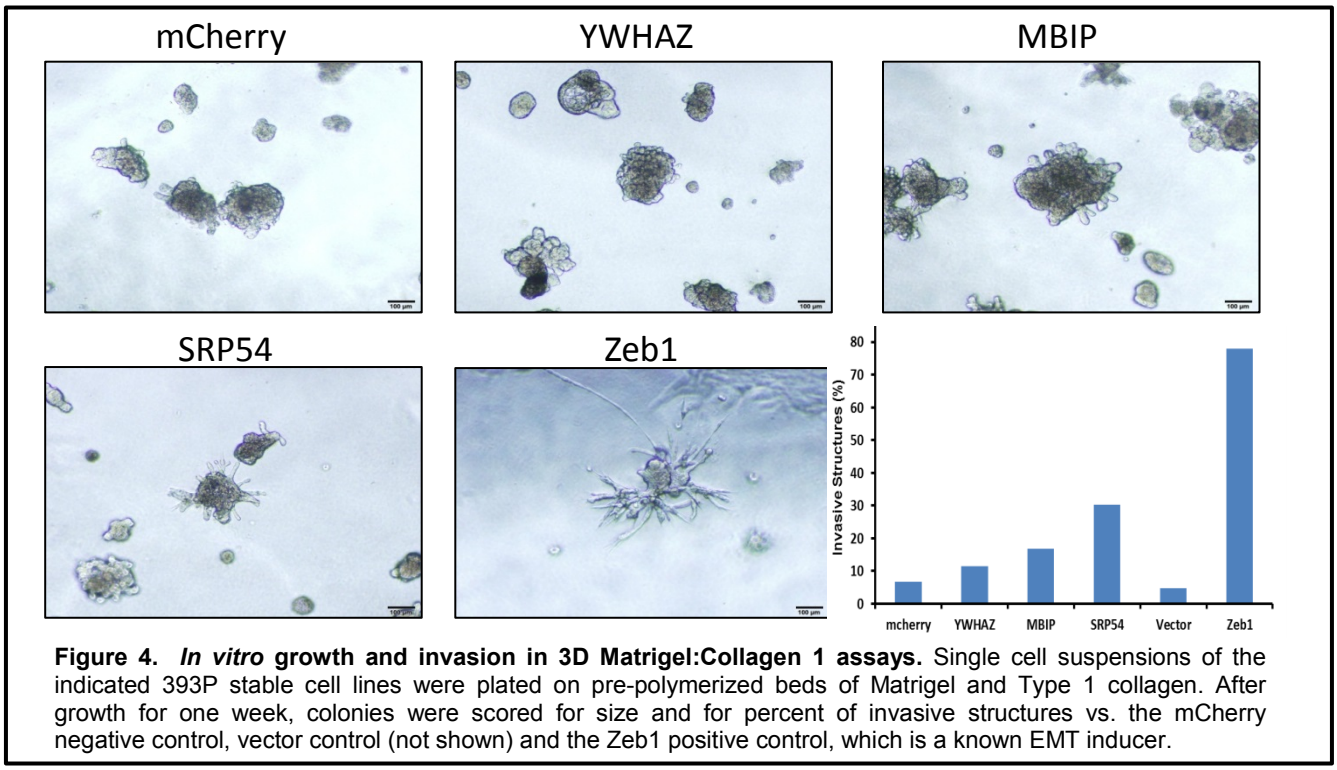

Figure 4. ***In vitro*** **growth and invasion in 3D Matrigel:Collagen 1 assays.** Single cell suspensions of the indicated 393P stable cell lines were plated on pre-polymerized beds of Matrigel and Type 1 collagen. After growth for one week, colonies were scored for size and for percent of invasive structures vs. the mCherry negative control, vector control (not shown) and the Zeb1 positive control, which is a known EMT inducer.

within 2 months (~12/01/2013). Given this delay, we chose to enter cells stably transduced with the ORF library virus into **3D culture assays,** whereby cells are grown on a bed of extracellular matrix composed of laminin-rich Matrigel or a mix of Matrigel and Type I collagen (Fig. 4). We have previously published that this method better mimics the *in vivo* condition by modeling the cell-matrix interactions, is scorable for cell growth, morphology and invasion, and can be modified by adjusting the composition and biophysical properties of the matrix [4-6]. As such we feel the 3D matrix assay is more robust and informative than the anoikis resistance assay, though we are continuing to develop the anoikis screens as described above. Our goal is to enter all ORF candidates into the 3D matrix assay given its tractability and robust output. Given our experience with this culture technique and the staining and microscopy in 3D cultures, this work is being conducted in the Gibbons lab.

Subaim 2.2. *In vivo* positive selection screens (proposed completion months 3-15).

In this Subaim we propose to perform *in vivo* metastasis screens with pooled viral-infected cells to positively select for single and combinatorial drivers of metastasis. Pilot assays demonstrated that subcutaneous injection of 10^6 cells provided an optimal period of primary tumor growth that could be monitored over ~6-8 weeks before the animals required sacrifice due to primary tumor burden (data not shown), thus allowing sufficient time for metastasis and detection of metastasis drivers. Although the cells can be inoculated orthotopically, the window of tumor growth in this case is reduced to only a few weeks because the animals develop labored breathing from the tumor burden in the lung, requiring earlier sacrifice. Since the goal of this *in vivo* screen is to recover distant metastases, the longer period of primary tumor growth is favored. After obtaining regulatory approval of the animal protocols from the MD Anderson IACUC and the DoD IACUC, cells infected and selected for stable expression of the ORFs in our library (as outlined above) were injected into the flank of syngeneic immunocompetent 129Sv mice. Based upon pilot experiments of optimal tumor cell inoculum, each group of injections included 10 mice that received a pool of 10^6 cells each, comprised of ORF expressing cell lines (50,000 cells/ORF) plus one additional mCherry control cell line used as an internal control. With each experimental pool we also injected 5 additional mice as negative controls with 10^6 cells expressing mCherry only. Mice are observed for general health and the growth of primary subcutaneous tumors recorded (representative cohort shown in Fig. 5A,B). The animals are sacrificed when the tumor burden is excessive or the mice show apparent deterioration in general health. This work has all been performed in the Gibbons lab and in the MDACC animal facility.

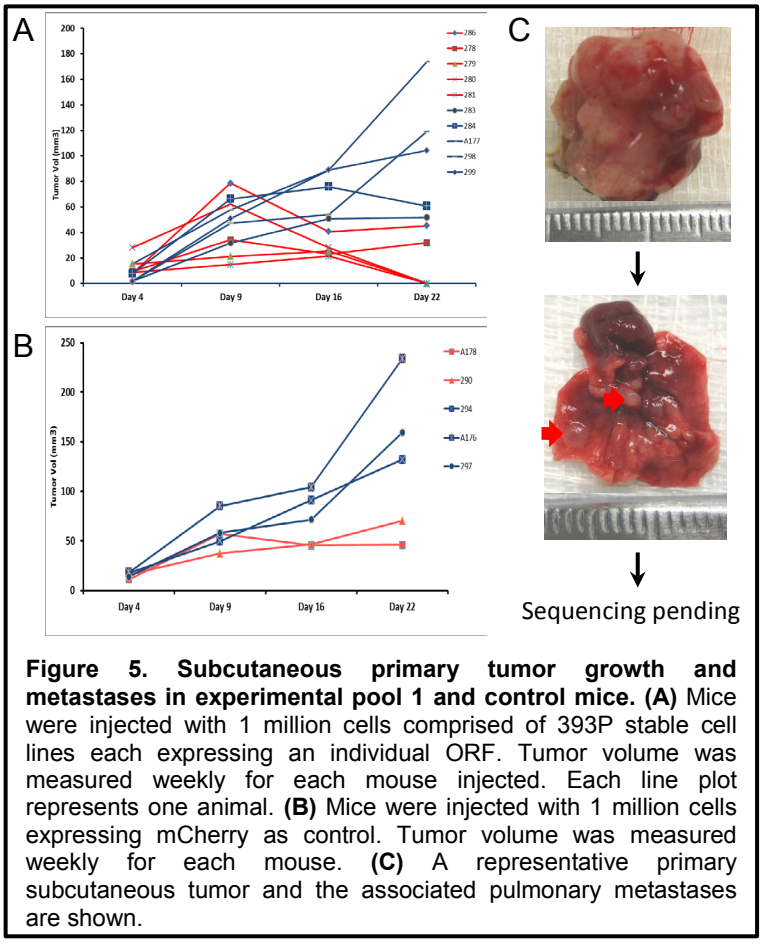

Figure 5. Subcutaneous primary tumor growth and metastases in experimental pool 1 and control mice. (A) Mice were injected with 1 million cells comprised of 393P stable cell lines each expressing an individual ORF. Tumor volume was measured weekly for each mouse injected. Each line plot represents one animal. **(B)** Mice were injected with 1 million cells expressing mCherry as control. Tumor volume was measured weekly for each mouse. **(C)** A representative primary subcutaneous tumor and the associated pulmonary metastases are shown.

As discussed above, we have opted for a "rolling" screening process whereby individual candidate pools of the 279 library panel are being entered into both the *in vitro* and *in vivo* screens

simultaneously to eliminate added costs and increase efficiency. We have entered approximately 75 of the 279 candidates into *in vivo* screens across multiple cohorts that are still under experimentation (i.e., tumors are actively growing and metastases pending). We are plotting survival and tumor growth curves (representative tumors shown in Fig. 5), and performing detailed animal necropsies, tissue collection and full pathological examination of all tissues in these animals (Fig. 5C shows a representative primary tumor and lung metastases). The tumors and identified metastases are frozen at the time of necropsy for subsequent DNA purification and sequencing of the unique barcodes, to identify by enrichment analysis the potential metastasis drivers in the pool (Fig. 6).

For potential drivers identified by *in vivo* selection and sequencing, validation of the *in vivo* results are being conducted by re-testing the individual cell populations (not in a pooled fashion) for metastasis formation *in vivo*. To date 2 identified potential drivers from both the *in vitro* and *in vivo* screens are being re-tested in animals.

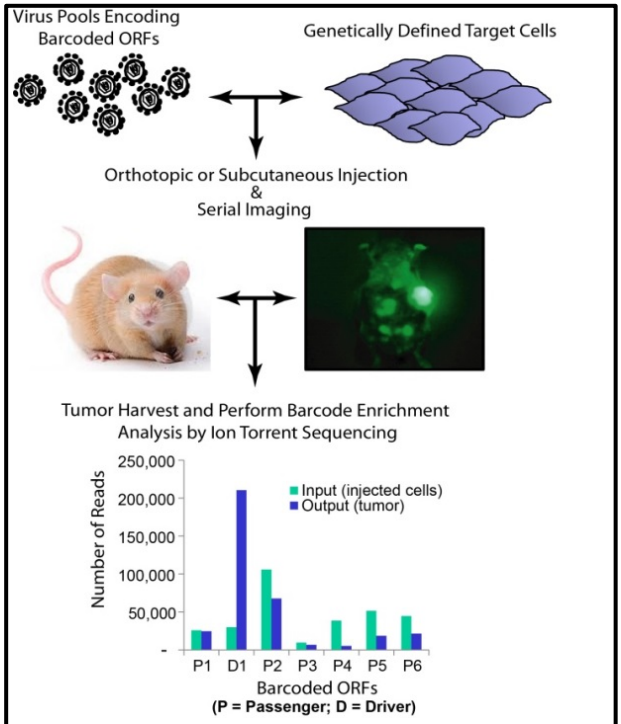

Figure 6. Enrichment analysis for drivers. ORFs are introduced into cells (input) for injection into animals. Primary tumors and metastases (output) are harvested, macrodissected for identification of the proviral insert (by barcodes). Histogram serves as data example for positive enrichment analysis of driver (D; *HoxA1*) over passengers (P).

Specific Aim 3: Clinicopathological prioritization and validation of top candidates.

In this Aim we propose to take forward identified drivers from the *in vitro* and *in vivo* screens for clinical validation and prioritization by analysis of their expression in publically available datasets (e.g., TCGA and others, as well as in-house datasets with clinical annotation). We also propose to begin functional and mechanistic studies of clinically validated high priority candidates.

Subaim 3.1. Clinical validation and prioritization of metastasis genes (proposed completion during Year 2).

This work is still pending the results of screening, but will proceed within the next 3-6 months.

Subaim 3.2. Functional and mechanistic study of the lead metastasis genes (proposed completion during Year 2).

For oncogenic and metastasis driver validation assays, we proposed to use cancer cell lines for appropriate expression or knock-down studies. Depending upon the particular genes identified in the screens and the observed phenotypes, we will use human NSCLC cancer cell lines or other cancer cell types as appropriate. In preparation for these mechanistic studies we have also developed a unique normal human bronchial epithelial cell (HBEC) line that will be particularly useful. These HBECs provide a lineage-specific cell line engineered with signature aberrations to dissect driver mechanism in the proper genetic context. Parental HBECs, which were obtained from Dr. John Minna (U.T. Southwestern), were immortalized by expression of hTERT and contain aberrant Cdk4 and knockdown of p53. We used this line to develop a doxycycline inducible $KRAS^{G12D}$ construct (pInducer-$KRAS^{G12D}$) based on our modified version of the pInducer vector system [7]. To test this cell model, virus carrying inducible GFP or oncogenic $KRAS^{G12D}$ was used to infect and select for stable HBEC cells, in which the expression of *KRAS* is controlled by treatment with or without doxycline (Dox; 100ng/ml) for 2 days. Immunoblotting for *KRAS* activation by RAF pull down assays indicated

significant *KRAS* activity in Dox-treated *KRAS*^{G12D} cells, whereas no activity was detected in the absence of Dox and in GFP cells (with or without Dox; Fig. 7 of Dr. Scott's Report). It is noteworthy that, compared to GFP and –DOX controls, Dox-induced HBEC-iKRAS^{G12D} cell lysates did not reveal significantly-elevated levels of total Kras protein owning to the low and titratable expression of the pInducer system. This suggests that our inducible model permits *KRAS* expression at a range near cell physiological levels that are desired for *KRAS* studies.

KEY RESEARCH ACCOMPLISHMENTS

- **Validation of a high-throughput *in vitro* invasion screens with standard larger format transwell assays:** Allows the rejection of false positives from the primary screens and identifies potential genes for further in vitro testing.

- **Robust 3D invasion assay:** Better mimics the 3D nature and extracellular matrix components found *in vivo*, while allowing manipulation of the matrix and easy scoring of the cellular phenotype.

- **Positive *in vivo* growth and metastasis screen:** Ability to directly test the *in vivo* role of genes on primary tumor growth and metastasis in a medium-throughput, pooled fashion due to the innovative barcoding techniques, use of sensitive sequencing and the positive selection of the *in vivo* screen.

REPORTABLE OUTCOMES

- **Pro-invasion genes:** Primary screens and secondary validation have revealed a number of genes that robustly drive cell invasion (e.g., *YWHAZ*, *MBIP* and *SRP54*). A full list of functionally annotated scoring genes among the 260 candidates screened will be provided in our next report and resulting publications.

CONCLUSIONS

Our overall goal with this project is to establish a pipeline of robust screening techniques to functionally prioritize the data emerging from large-scale genomics efforts in lung cancer. Using a combination of *in vitro* and *in vivo* screens we will be able to identify and validate oncogene and metastasis drivers, explore the mechanistic basis for their function, and generate the pre-clinical cell and animal models needed for therapeutic targeting studies. As outlined in the Body of the main report from Dr. Scott, we have successfully constructed the necessary libraries for this work, implemented both the *in vitro* and *in vivo* screens, and are putting into place the necessary tools for mechanistic studies of the lead candidate genes. We have made several technical improvements to our overall work plan and view this first year as a robust start to further screening of the TCGA data (which has only recently been released) that will extend past the life of this particular grant. We also feel that these techniques can be broadly implemented for functionalization of genomic data for other tumor types, e.g. pancreatic cancer.

REFERENCES

1. Scott, K.L., et al., *Proinvasion metastasis drivers in early-stage melanoma are oncogenes.* Cancer Cell, 2011. **20**(1): p. 92-103.

2. Wardwell-Ozgo, J., et al., *HOXA1 drives melanoma tumor growth and metastasis and elicits an invasion gene expression signature that prognosticates clinical outcome.* . Oncogene, 2013. **Epup ahead of print**.
3. Douma, S., et al., *Suppression of anoikis and induction of metastasis by the neurotrophic receptor TrkB.* Nature, 2004. **430**(7003): p. 1034-9.
4. Gibbons, D.L., et al., *Contextual extracellular cues promote tumor cell EMT and metastasis by regulating miR-200 family expression.* Genes Dev, 2009. **23**(18): p. 2140-51.
5. Ahn, Y.H., et al., *ZEB1 drives prometastatic actin cytoskeletal remodeling by downregulating miR-34a expression.* J Clin Invest, 2012. **122**(9): p. 3170-83.
6. Gill, B.J., et al., *A synthetic matrix with independently tunable biochemistry and mechanical properties to study epithelial morphogenesis and EMT in a lung adenocarcinoma model.* Cancer Res, 2012. **72**(22): p. 6013-23.
7. Meerbrey, K.L., et al., *The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo.* Proc Natl Acad Sci U S A, 2011. **108**(9): p. 3665-70.

APPENDICES

None